Crime and punishment

Weissing, Franz; Ostrom, Elinor

Published in:
Journal of Theoretical Politics

DOI:
10.1177/0951692891003003006

Crime and Punishment: Further Reflections on the Counterintuitive Results of Mixed Equilibria Games

Franz Weissing and Elinor Ostrom

Journal of Theoretical Politics 1991; 3; 343

DOI: 10.1177/0951692891003003006

Journal of Theoretical Politics 3(3): 343-350 (1991)

RESEARCH NOTE

CRIME AND PUNISHMENT: FURTHER REFLECTIONS ON THE COUNTERINTUITIVE RESULTS OF MIXED EQUILIBRIA GAMES

Franz Weissing and Elinor Ostrom

ABSTRACT

In a series of related articles, George Tsebelis (1989, 1990, 1991) challenges political theorists to rethink the foundations of policy analysis. His major critique of policy analyses, based on decision theory (where one individual decides in an inanimate but not certain environment) rather than on game theory (where one individual decides in an environment with other strategic individuals), has weathered the storm of commentaries made on his work. Tsebelis's argument, that payoff changes for one player do not affect the behavior of that player at a mixed-strategy equilibrium, holds in some cases but not in others. Whether changes in the payoffs of one player affect that player's behavior at a mixed-strategy equilibrium depends upon the type of linkages that exist between that player and the others in the frame or in a similar position in a game. In this note we have stated the general conditions under which changes in the payoffs of one player do not affect that player's behavior at equilibrium and when they do.

KEY WORDS • Nash equilibrium • regulation game • strategic interdependence

In a series of related articles, George Tsebelis (1989, 1990, 1991) challenges political theorists to rethink the foundations of policy analysis. His major critique of policy analyses based on *decision* theory (where one individual decides in an inanimate but not certain environment) rather than on *game* theory (where one individual decides in an environment with other strategic individuals), has weathered the storm of commentaries made on his work (Bianco et al., 1990; Hirshleifer and Rasmusen, 1990; Mayer, 1991; Rapoport, 1990). Illustrating that many important policy problems are better cast as strategic interaction problems rather than as simple probabilistic choice is an important contribution.

Tsebelis's second major argument within the game theory tradition has not fared as well. Tsebelis concludes that in a 2-person, strategic interaction problem – represented as a game without pure strategy equilibria – changes in the payoffs of the first player influence the behavior of the second player, but do not influence the first player's behaviour. For example, Tsebelis (1991) argues in this journal that the equilibrium behavior of firms in a regulated industry is not affected by changes in penalties assessed against firms who break the law. In a similar vein, he asserts that penalties assessed against those who break a law do not affect the equilibrium behavior of the public, and that economic sanctions imposed by international bodies do not affect the equilibrium behavior of errant nations (Tsebelis, 1989). These counterintuitive aspects of mixed-strategy equilibria have been noticed by other scholars (Wittman, 1985; Holler, 1990). The sweeping conclusions that Tsebelis draws, however, from the 2-person example, have aroused considerable attention and critique.

Tsebelis claims that n-person games are likely to have similar counterintuitive results to those of 2-person games. Critics are skeptical that this problem will be found in anything other than 2-person games. Weissing and Ostrom (1991) show that the behavior of some players in a $n > 2$ person game at equilibrium is affected only by changes in other players' payoffs – as Tsebelis predicted. We do *not* find, however, that this is *generally* the case in the $n > 2$ person setting. Further, Tsebelis's claim that this always occurs in 2-person games is challenged by our findings. Since these are important issues at the core of modern political theory, we have mused further about the relationship of individual payoffs, beliefs about other players' strategies, and changes in behavior at equilibrium.

In an insightful critique, Ordeshook (in Bianco et al., 1990: 573) stresses that in games with $n > 2$ players:

> . . . an equation stating the mixed equilibrium solutions for one player must be formulated in terms of the remaining $n-1$ players' strategies (thereby formalizing the game-theoretic reasoning that, conditional on beliefs, each player chooses a strategy that is a best response to the strategies of everyone else) . . . .

Ordeshook's reminder that Nash equilibria are based on the beliefs held by each player about the strategies of all others suggests a way of representing mixed, Nash equilibria that clarifies several issues.

The Unified Actor Case

Tsebelis focuses on single actors representing classes of actors. One unified policeman represents all police. One unified player represents all of the public. This is how he addresses broad policy questions with a 2-person game. In this section, we retain the unified actor assumption.

Let us first posit a 2-person game where each player (i.e. each unified actor) has two pure strategies. The players are denoted by A and B, their pure strategies by **1** and **2**, and mixed strategies by s_A and s_B. The payoff expected by player A for each of his two pure strategies is a function E_A involving player A's own payoff parameters and player A's beliefs concerning the strategic behavior of player B (denoted by $[s_B]_A$):

$$E_A\ (\mathbf{1} \mid [s_B]_A) \tag{1a}$$

$$E_A\ (\mathbf{2} \mid [s_B]_A) \tag{1b}$$

If player A is rational, he would use that strategy which yields him the highest expected payoff. In particular, player A would only use a completely mixed strategy if he expects to get the same payoff for his own pure strategies:

$$\Delta E_A\ ([s_B]_A) = 0 \tag{2}$$

where ΔE_A denotes the difference between (1a) and (1b).

Equation (2) illustrates that the strategic considerations of player A are affected only by the payoff parameters of player A and by his belief about the strategic choice of the other player. In the 2-person context, these two components can be decoupled. If ΔE_A is invertible, (2) can be rewritten in the form:

$$[s_B]_A = (\Delta E_A)^{-1}(0) \tag{3}$$

Equation (3) makes it obvious that player A's belief concerning the strategic choice of player B is a function of the payoff parameters of player A. Intuitively speaking, player A will only use a completely mixed strategy if he believes that player B makes his strategic decision dependent on player A's behavior (which, in turn, is governed by player A's payoff parameters). It is a characteristic property of a Nash equilibrium that no player can get an advantage by deviating unilaterally from the strategy which is expected by the other players. Therefore, the beliefs of the players are consistent with the actual equilibrium behavior (denoted by s_A^* and s_B^*):

$$s_A^* = [s_A]_B \tag{4a}$$

$$s_B^* = [s_B]_A \tag{4b}$$

Combining (3) and (4) (i.e. payoff equivalence of pure strategies at a mixed-strategy equilibrium and consistency of beliefs and behavior), we get Tsebelis's main result: at equilibrium, the strategic choice of player B is a function of the payoff parameters of player A if the equilibrium strategy of A is a completely mixed strategy.

Let us now investigate how far this result can be generalized to the n-person context. For illustrative purposes, we concentrate on the 3-person case. The three players are denoted by A, B and C, their two pure strategies by **1** and **2**, and their mixed strategies by s_A, s_B and s_C. As in the 2-person context, player A's expected payoff for his pure strategies is given by a function E_A involving player A's payoff parameters and player A's beliefs (denoted by $[s_B]_A$ and $[s_C]_A$) concerning the strategic behavior of the other players. However, player A will take into account only those players who actually have a strategic influence on player A's payoff.

Player A will use a completely mixed strategy only if he expects to get the same payoff from his two pure strategies:

$$\Delta E_A([s_B]_A, [s_C]_A) = 0 \tag{5}$$

At a Nash equilibrium, beliefs are compatible with equilibrium behavior:

$$s_A^* = [s_A]_B = [s_A]_C \tag{6a}$$

$$s_B^* = [s_B]_A = [s_B]_C \tag{6b}$$

$$s_C^* = [s_C]_A = [s_C]_B \tag{6c}$$

Taken together, (5) and (6) give the following conditions for a Nash equilibrium in completely mixed strategies:

$$\Delta E_A(s_B^*, s_C^*) = 0 \tag{7a}$$

$$\Delta E_B(s_A^*, s_C^*) = 0 \tag{7b}$$

$$\Delta E_C \, (s_A{}^*, s_B{}^*) = 0 \tag{7c}$$

If the expected payoff of one player, say A, is affected only by the strategic choice of one other player, say B, (7a) is essentially equivalent to $\Delta E_A \, (s_B{}^*) = 0$ which implies that player B's equilibrium strategy may be viewed as a function of player A's payoff parameters. Correspondingly, Tsebelis's result that the equilibrium strategy of a player B depends only on the payoff parameters of one other player A holds true if:

- player A uses a completely mixed strategy; and
- player A's expected payoff is affected only by the behavior of player B and not by the behavior of the other players.

In general, however, the equilibrium strategy of player B is indirectly dependent on his own payoff parameters. In fact, $s_B{}^*$ is linked to $s_C{}^*$ via (7a), whereas $s_C{}^*$ is linked to the payoff parameters of player B via (7b). This feedback loop between the equilibrium strategy of a player and his payoff parameters makes Ordeshook's argument somewhat more transparent. It is also possible to generalize to cases where Tsebelis's argument is nevertheless true.

When the feedback loops are cut, Tsebelis's conclusion holds. Now, how might the feedback loops be severed? Let us assume that two players interact and affect one another, but each is affected by different sets of other players. This implies that player A's equilibrium strategy can be viewed as a function of player B's payoff parameters and the equilibrium strategies of players C and D. Players C and D affect player B's payoff, but not player A's expected payoff. The behavior of C and D is not connected to the payoff parameters of player A. Thus, the feedback loop is cut. When this occurs, player A's equilibrium behavior can be viewed as a function of the other player's payoff parameters but not his own.

One additional condition is required in order to cut the feedback loop. Let us assume that there are two disjunct sets of players, G_1 and G_2, which are 'complements' in the following sense: whenever a player is affected by one or more players in G_1, he is not affected by a player in G_2 and vice versa. Let us now assume that players A and B interact and affect one another, but that A is affected by players in G_1 whereas B is affected by players in G_2. Then, at an equilibrium where A and B both use a mixed strategy, player A's and B's equilibrium behavior can be viewed as a function of other players' payoff parameters but not their own.

Types of Players Rather than Unified Players

Tsebelis looks at players (police and the public; regulators and regulated firms) as unified players. In real-world settings, these players may or may not be unified. What happens if all the players treated as unified players above, A, B and C, are no longer viewed as unified players? Rather, we now view A, B and C as positions (or as a certain type of player).[1] Position A, say a policeman, is now occupied by multiple players who are treated completely symmetrically. Position B, say a citizen, is also occupied by multiple symmetric players. By symmetry, we mean:

1. There is no clear notational convention within game theory to denote the difference between players and positions while in political theory, the distinction is quite important.

1. All players in a position are affected equally by events; this is reflected by assigning an identical set of payoff parameters to each player in a position.
2. All players are structurally equivalent in that similar behavior leads to similar consequences for the other players of their type.
3. All players come from the same population with a prevailing norm of behavior, and behave similarly whenever they find themselves in a similar situation.
4. All players have the same reliable knowledge about the game.
5. All players in a position face identical starting conditions.

If one assumes that each of the players in position A is not affected by the other players in position A, but only by the players in the other positions, then all of the results in the first section remain true. For example, if we take the Tsebelis game between the police and the public, then each policeman would be affected only by the behavior of players in the position of citizen, and not by other policemen. Each citizen is affected only by the behavior of players who are policemen, and not by the behavior of other citizens. In this case, the Tsebelis results remain true. The rate of crime at a mixed-strategy equilibrium depends on the incentives of the police and not on the incentives of the public.

In many situations, however, the incentives of the players in the same position are linked in some manner. We rely on the irrigation game that we have analyzed fully elsewhere (Weissing and Ostrom, 1991). Let us imagine a population of farmers who share a single irrigation system that is their main source of water for their crops. The right to take water is rotated among the farmers according to a predetermined formula. At each point of time, only one farmer is in the position of a turntaker (TT) who is authorized to take a certain quantity of water. The other farmers wait for their turn to take water and are called turnwaiters (TWs). A potential conflict of interest exists between the turntaker and the turnwaiters, because the turntaker could extract more than the authorized amount of water from the system. A player in the position of the turntaker usually has an incentive to steal. The turnwaiters are all negatively affected if the turntaker steals. At any time, only one person is in the position of TT and chooses to steal or not; *n* players are in the position of a TW who can monitor or not.

In this game, the players in the position of the TW are mainly concerned with deterring or preventing stealing. Who detects a stealing event is somewhat important for them, but their main concern is whether the stealing event is detected at all. Therefore, the expected payoff of a TW depends not only on the stealing rate but also on the probability that stealing is detected by other TWs. Even if all the TWs are treated completely symmetrically, our general analysis shows that at a mixed-strategy equilibrium, the behavior of the TT depends not only on the payoff parameters of the TW but also on the payoff parameters of the TT himself.

In order to explain this result, let us now concentrate on the interaction between players who could be in one of two positions, A or B. First assume that only one player is in position A (the TT), whereas *n* players are in position B (the TW). Let us ask, what are the strategic considerations of the player in position A? In the symmetric context, the equilibrium payoff for each of his two pure strategies is a function of his own payoff parameters and his beliefs concerning the strategies of the other players. The important point is that in the symmetric context, player A should believe that all the other players in position B will play the same strategy. In other words, his beliefs should be symmetric. Therefore, his equilibrium considerations should be governed by his 'unified' beliefs about the other players.

$$E_A(1|\ [s_B]_A) = E_A(2|\ [s_B]_A) \tag{8}$$

At a symmetric equilibrium, beliefs should be consistent with the behavior. Or,

$$s_B^* = [s_B]_A \tag{9}$$

Therefore (8) yields the following equilibrium condition:

$$\Delta E_A(s_B^*) = 0 \tag{10}$$

Thus s_B^* is a function of player A's payoff parameters provided that player A uses a mixed strategy. This explains why the TW behavior can be viewed as being dependent only on the TT payoff parameters.

Looking at the strategic consideration of a player in position B, the situation is different. The expected payoffs for each of the two pure strategies of a player in position B, say X, are dependent not only upon his beliefs concerning the behavior of the players in position A, but also on the behavior of the other players in his own position B.

$$E_B(1|\ [s_A]_X, [s_B]_X) = E_B(2|\ [s_A]_X, [s_B]_X) \tag{11}$$

At a symmetric equilibrium in completely mixed strategies, the beliefs of a specific player, X, have to be consistent with the actual behavior of the other players:

$$s_A^* = [s_A]_X \text{ and } s_B^* = [s_B]_X$$
$$\text{for all X in position B} \tag{12}$$

Therefore, at equilibrium:

$$\Delta E_B(s_A^*, s_B^*) = 0 \tag{13}$$

Equation 13 implicitly gives the equilibrium behavior of player A as a function of payoff parameters in position B and the equilibrium strategies of the players in position B. On the other hand, we have already seen in (10) that s_B^* is a function of payoff parameters in position A. This implicitly shows that player A's equilibrium strategy depends on both the payoff parameters in position A and in position B.

If there are several players in position A that affect each others' payoffs and behavior, then Tsebelis's argument breaks down completely. Equilibrium strategies of both types of players depend on the payoff parameters of both types of players. Whenever it makes better sense to model a position, such as a turnwaiter in an irrigation system, as a number of identical players whose behavior affects each other's payoffs, rather than as a unified agent, the argument that Tsebelis makes concerning games with $n > 2$ players does not apply.

Conclusion

The important point that Tsebelis makes is that strategic choices differ much from parametric choices. Most policies involving police and the public, or regulators and regulated firms, involve strategic choices rather than parametric choices. This point should not be lost in the debate. Tsebelis's argument, that payoff changes for one player do not affect the behavior of that player at a mixed-strategy equilibrium, holds in some cases and does not hold in others. Whether changes in the payoffs of one player affect that player's behavior at a mixed-strategy equilibrium depends upon the type of linkages that exist between that player and the others in the game or in a similar position in a game. In this note, we have stated the general conditions when changes in the payoffs of one player do not affect that player's behavior at equilibrium and when they do. It is now clear that there are many games involving mixed-strategy equilibria where a change in one player's payoff parameters does affect that player's behavior at equilibrium. If is also clear that this is not always the case. How players are linked to one another is a key factor affecting this relationship.

REFERENCES

Bianco, W.T., P. Ordeshook and G. Tsebelis (1990) 'Crime and Punishment: Are One-Shot, Two-Person Games Enough?', *American Political Science Review* 84: 569–86.

Hirshleifer, J. and E. Rasmusen (1990) 'Are Equilibrium Strategies Unaffected by Incentives? The Police Game'. Los Angeles: University of California, Department of Economics, Working Paper 595.

Holler, M. J. (1990) 'The Unprofitability of Mixed-Strategy Equilibria in Two-Person Games', *Economics Letters* 32: 319–23.

Mayer, T. F. (1991) 'Comments on Tsebelis's "Penalty Has No Impact on Crime"', *Rationality and Society* 3: 266–8.

Rapoport, A. (1990) 'Comments on Tsebelis', *Rationality and Society* 2: 508–11.

Tsebelis, G. (1989) 'The Abuse of Probability in Political Analysis: The Robinson Crusoe Fallacy', *American Political Science Review* 83: 77–92.

Tsebelis, G. (1990) 'Penalty Has No Impact on Crime: A Game-Theoretical Analysis', *Rationality and Society* 2: 255–86.

Tsebelis, G. (1991) 'The Effect of Fines on Regulated Industries: Game Theory vs. Decision Theory', *Journal of Theoretical Politics* 3(1): 81–101.

Weissing, F. and E. Ostrom (1991) 'Irrigation Institutions and the Games Irrigators Play: Rule Enforcement without Guards', in R. Selten (ed.) *Game Equilibrium Models II: Methods, Morals, and Markets*. Berlin: Springer-Verlag, forthcoming.

Wittman, D. (1985) 'Counter-Intuitive Results in Game Theory', *European Journal of Political Economy* 1: 77–89.

FRANZ WEISSING is a theoretical biologist at the University of Groningen, Netherlands. As a theoretical population biologist, he is interested in evolutionary games that describe individual behavior in ecological settings. He is author of 'Evolutionary Stability and Dynamic Stability in a Class of Evolutionary Normal Form Games' (in R. Selten, ed., *Game Equilibrium Models I: Evolution and Game Dynamics*, pp. 29–97, Springer-Verlag, Berlin).

ELINOR OSTROM is Co-Director of the Workshop in Political Theory and Policy Analysis and the Arthur F. Bentley Professor of Government at the Department of Political Science at Indiana University, Bloomington. She is the author of *Governing the Commons*, co-author with Robert Bish and Vincent Ostrom of *Local Government in the United States*, and co-editor with Richard Kimber and Jan-Erik Lane of *Journal of Theoretical Politics*.